UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/058,143 | 01/29/2002 | Sara H. Basson | YOR920010346 | 7380 |

7590 02/09/2005

McGuireWoods LLP
Suite 1800
1750 Tysons Boulevard
McLean, VA 22102-3915

| EXAMINER | |
|---|---|
| ALBERTALLI, BRIAN LOUIS | |
| **ART UNIT** | **PAPER NUMBER** |
| 2655 | |

DATE MAILED: 02/09/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 10/058,143 | BASSON ET AL. | |
| | Examiner | Art Unit | |
| | Brian L Albertalli | 2655 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) 1-20 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-20 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 06 September 2002 is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

## DETAILED ACTION

### *Specification*

1. The summary of the invention is objected to because it is not completed. Applicant is reminded that the summary of the invention must be completed in order for the application to be in condition for allowance.

### *Drawings*

The drawings are objected to because:

a) In Fig. 2, element 212, "SENDOR" should be –SENDER--.

b) In Fig. 6., element 601, "IS" should be –DOES--.

Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. The objection to the drawings will not be held in abeyance.

### *Claim Objections*

2. Claim 13 is objected to because of the following informalities: in line 2 of the claim, the quotation mark after the word "credibility" should be removed. Appropriate correction is required.

3. Claims 2, 5, 8, and 20 are objected to because the use of the term "analyzed voice data" to refer to the data analyzed associated with a first voice is ambiguous. For example, in line 11 of claim 2, the phrase "integrating the analyzed voice data and the

analyzed second voice data" is confusing because "the analyzed voice data" could refer to the analyzed data associated with the first voice or the second voice. The Examiner suggests that claims 2, 5, 8, and 20 be amended so that all occurrences of the term "the analyzed voice data" read –the analyzed first voice data--.

### *Claim Rejections - 35 USC § 102*

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

5. Claims 14 and 15 are rejected under 35 U.S.C. 102(e) as being anticipated by Walker (U.S. Patent Application Publication 2003/0050777).

In regard to claim 14, Walker discloses a system for integrating acoustic data using speech recognition (Fig. 1, 10), comprising:

a communication module (server 38) which receives voice data from a plurality of computers each having speech recognition residing thereon (personal computers 16, 22, 28, and 34 with respective speech recognizers 12, 18, 24, and 30), the communication module residing on the plurality of computers or a remote server (page 1, paragraph 11, lines 1-5 and page 2, paragraph 12)

an evaluator module associated with each of the plurality of computers, the evaluator module analyzes the voice data from each of the plurality of computers (transcription service 36 analyzes the for user information and time stamps of each entry, page 1, paragraph 11, lines 6-9); and

an integrator module associated with the evaluator module, the integrator module integrates all of the analyzed voice data from each of the plurality of computers and provides one decoding output (the transcription entries are arranged as an ordered and interleaved transcription of the plurality of computers, page 1, paragraph 11, line 9 through page 2, 1st column, line 3).

In regard to claim 15, Walker discloses:

the voice data is associated with at least two master speakers associated with the speech recognition associated with different computers of the plurality of computers (speech recognizers 12, 18, 24, and 30 are associated with particular first through fourth persons, respectively, page 1, paragraph 8, lines 9-12 and paragraph 10, lines 4-14); and

the integrator module integrates the voice data of the at least two master speakers into the one decoding output (the conversation of several persons is interleaved by transcription service 36, page 2, 1st column, lines 1-3 and paragraph 13, lines 3-11).

***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

7. Claims 1-13, 16-17, and 19-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Walker, in view of Bennett et al. (U.S. Patent 6,701,293).

In regard to claims 1 and 19, Walker discloses a method and a machine readable medium containing code for integrating acoustic data using speech recognition (Fig. 2), comprising the steps of:

detecting voice data on a first computer (voice data from a first user is detected by first speech recognizer on a first computer, step 100, page 2, paragraph 15, lines 4-7);

identifying the voice data as a first master speaker associated with a speech recognition system residing on the first computer (step 100, the speech recognizer recognizes the voice data as coming from person #1, page 2, paragraph 15, lines 4-7);

analyzing the voice data residing on the first computer (step 102, the speech recognizer converts the utterance into a dictation including text, page 2, paragraph 15, lines 7-11); and

integrating the analyzed voice data from the first computer into a single decoding output (step 106, results from a plurality of speech recognizers are integrated into an interleaved transcript, page 2, 2nd column, lines 6-9).

Walker further discloses detecting a second voice data on a second computer and analyzing and integrating that second voice data into the output transcript (see Fig. 2).

Walker does not disclose providing the first voice data from the first computer to the at least second computer having a speech recognition system thereon and recognizing the first voice data in parallel on both the first and at least second computers.

Bennett et al. disclose a method for recognizing speech (Fig. 3) that analyzes voice data (input stream) on a plurality of speech recognizers and integrates the results of those recognizers into a single output (column 3, lines 51-53, and column 4, lines 20-21).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Walker to provide a first voice data to each of the speech recognizers 12, 18, 24, and 30 in parallel, and then combine the results of those recognizers into a single decoding output, as taught by Bennet et al., in order to increase the recognition accuracy of the overall system and provide a more accurate transcript of the voice data.

In regard to claims 2 and 20, Walker discloses:

detecting a second voice data on at least the second computer (voice data from a second user is detected by a second speech recognizer on a second computer, page 2, paragraph 16, lines 3-7);

identifying the second voice data as being a second master speaker associated with the speech recognition system of the at least the second computer (step 108, voice data is recognized as coming from person #2, page 2, paragraph 16, lines 3-7);

analyzing the second voice data residing on the at least the second computer (step 110, the speech recognizer converts the utterance into a dictation including text, page 2, paragraph 16, lines 7-10); and

integrating the analyzed voice data and the analyzed second voice data into the single decoding output (step 106, results from a plurality of speech recognizers are integrated into an interleaved transcript, page 2, paragraph 26, lines 12-17).

Walker does not disclose providing the second voice data from the at least second computer to the first computer and recognizing the first voice data in parallel on both the first and at least second computers.

Bennett et al. disclose a method for recognizing speech (Fig. 3) that analyzes voice data (input stream) on a plurality of speech recognizers and integrates the results of those recognizers into a single output (column 3, lines 51-53, and column 4, lines 20-21).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Walker to provide a second voice data to each of the speech recognizers 12, 18, 24, and 30 in parallel, and then combine the results of those recognizers into a single decoding output, as taught by Bennet et al., in order to increase the recognition accuracy of the overall system and provide a more accurate transcript of the voice data.

In regard to claim 3, Walker discloses the at least second computer is a second and third computer (Fig. 1, first computer 16, second computer 22, and third computer 38).

In regard to claim 4, Walker discloses each computer is associated with a person (speech recognizers 12, 18, 24, and 30 are associated with particular first through fourth persons, respectively, page 1, paragraph 8, lines 9-12 and paragraph 10, lines 4-14). Furthermore, Walker discloses that each recognizer identifies when its particular person is speaking (see, for example page 2, paragraph 15, lines 4-7). In Fig. 1, the audio associated with each person is shown closest to the respectively associated speech recognizer.

Walker is silent as to the details of how the first and second speakers are identified.

Official notice is taken that it is notoriously well known and recognized in the art that the volume of sound input received by a microphone increases dramatically when somebody is speaking into it.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Walker et al. to perform the step of detecting who was speaking by determining when a volume was higher than a predetermined background noise level threshold on that user's respective computer, because determining who was speaking

with a simple volume threshold comparison would greatly reduce the amount of processing needed to determine who was speaking.

In regard to claim 5, Walker discloses:

summarizing the analyzed voice data and the second voice data into a single transcript (the transcription entries are arranged as an ordered and interleaved transcription of the plurality of computers, page 1, paragraph 11, line 9 through page 2, 1st column, line 3).

In regard to claims 6 and 7, Walker does not disclose the analyzed voice data is weighted.

Bennet et al. disclose weighting the analyzed voice data differently for each of the plurality of recognizers (column 6, lines 7-16). The higher weight of the analyzed voice data would necessarily be selected as the most accurate rendition of the voice data.

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify Walker to weight the analyses of the first voice data and second voice data according to a weight assigned to the first recognizer and a weight assigned to the second recognizer, so that poor recognition results from one of the recognizers would not overly influence the final recognition result, thereby increasing the accuracy of the overall system.

In regard to claim 8, Walker does not disclose providing a confidence level for each word.

Bennett et al. disclose providing a confidence level for each recognition result generated by each recognizer (column 5, lines 6-8). The confidence level is used in the integration of the results of each recognizer (column 5, lines 10-14).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify Walker et al. to provide a confidence level to each word associated with both the first voice data and second voice data, so that at the integration step, the recognizer with the highest confidence level would be selected as the correct result, as taught by Bennet et al. This would ensure that the best recognition result was included in the transcript, thereby increasing the accuracy of the transcript and reducing the need for later editing of the transcript.

In regard to claim 9, the connections disclosed by Walker between the first computer and at least second computer (arrows in Fig. 1) must necessarily communicate in a wire or wireless communication protocol.

In regard to claim 10, Walker discloses:

the speech recognition of the first computer and the at least the second computer are one of (i) a same speech recognition system and (ii) a different speech recognition system (any number of speech recognition software applications are used for speech recognizers 12, 18, 24, and 30, page 2, paragraph 14, lines 1-3); and

the first master speaker and the second master speaker are further associated with the speech recognition of the at least the second computer and the first computer, respectively (speech recognizers 12, 18, 24, and 30 are associated with particular first through fourth persons, respectively, page 1, paragraph 8, lines 9-12 and paragraph 10, lines 4-14).

In regard to claim 11, neither Walker nor Bennett et al. specifically disclose filtering out the background noise.

Official notice is taken that it is notoriously well known and recognized in the art to filter out background noise, since background noise included in the speech data drastically reduces a speech recognizer's ability to accurately recognize speech.

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Walker and Bennett et al. to filter out background noise, since background noise included in the speech data drastically reduces a speech recognizer's ability to accurately recognize speech.

In regard to claim 12, neither Walker nor Bennett et al. disclose providing feedback to the first computer and at least second computer relating to the performance of the recognizers or maintaining a record of credibility.

Official notice is taken that it is notoriously well known and recognized in the art to provide feedback relating to the performance analysis of a speech recognizer so that the performance recognizer could be improved. Furthermore, it is notoriously well

known and recognized in the art to maintain a record of credibility relating to the ability to recognize a master speaker.

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Walker and Bennet et al. to track the speech recognition performance (the accuracy of determining *what* was being said) as well as the voice recognition performance (the accuracy of determining *who* was talking) so that both could be continuously adapted to increase the accuracy of both using any well known adaptation method.

In regard to claim 16, Walker does not disclose providing a confidence level for each word.

Bennett et al. disclose providing a confidence level for each recognition result generated by each recognizer (column 5, lines 6-8). The confidence level is used in the integration of the results of each recognizer (column 5, lines 10-14).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify Walker et al. to provide a confidence level to each word associated with both the first voice data and second voice data, so that at the integration step, the recognizer with the highest confidence level would be selected as the correct result, as taught by Bennet et al. This would ensure that the best recognition result was included in the transcript, thereby increasing the accuracy of the transcript.

In regard to claim 17, Walker discloses each computer is associated with a person (speech recognizers 12, 18, 24, and 30 are associated with particular first through fourth persons, respectively, page 1, paragraph 8, lines 9-12 and paragraph 10, lines 4-14). Furthermore, Walker discloses that each recognizer identifies when its particular person is speaking (see, for example page 2, paragraph 15, lines 4-7). In Fig. 1, the audio associated with each person is shown closest to the respectively associated speech recognizer.

Walker is silent as to the details of how the first and second speakers are identified.

Official notice is taken that it is notoriously well known and recognized in the art that the volume of sound input received by a microphone increases dramatically when somebody is speaking into it.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Walker et al. to perform the step of detecting who was speaking by determining when a volume was higher than a predetermined background noise level threshold on that user's respective computer, because determining who was speaking with a simple volume threshold comparison would greatly reduce the amount of processing needed to determine who was speaking.

Claim 18 is rejected under 35 U.S.C. 103(a) as being unpatentable over Walker, in view of Saito et al. (European Patent Application 1 061 724).

Walker discloses:

a final decoder output module associated with the integrator module, the final decoder output module prepares a summary of the decoded output; and

summurator module for receiving the summary of the decoded output (the conversation of several persons is interleaved and stored by transcription service 36, page 2, 1st column, lines 1-3 and paragraph 13, lines 3-11).

Walker does not disclose a sender module for sending the decoded output to a computer of the plurality of computers for transcription or editing the decoded output.

Saito et al. disclose a sender module for sending the decoded output to a computer of the plurality of computers for transcription or editing the decoded output (image display means 15, correction input mean 16, and correction executing means 17 function to send the decoded output to users for correction and editing, page 10, paragraph 52 and page 11, paragraphs 53 and 54).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Walker to include a sender module so that any misrecognized words could be corrected at the consent of all the users who's voice was transcribed, as taught by Saito et al. (column 14, lines 6-10).

***Conclusion***

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Fiscus (*A Post-Processing System to Yield Reduced Word Error Rates*) discloses a "Rover" voting system for multiple speech recognizers. Barry et al.

(*The Simultaneous use of Three Machine Speech Recognition Systems to Increase Recognition Accuracy*) discloses multiple speech recognizers operating on one input provide better recognition results. Din (U.S. Patent 6,754,631) discloses a method for transcribing meeting minutes. Joost (U.S. Patent 6,327,568) discloses a system that dynamically assigns speech recognition tasks to computers on a network. Schrage (U.S. Patent 6,850,609) discloses a device for creating transcripts of multiple speakers. Gudorf et al. (U.S. Patent 6,687,671) disclose a method for summarizing and distributing meeting transcripts. Ortega et al. (U.S. Patent 6,535,848) disclose a method for integrating multiple speakers voice data into one transcript. Chandler et al. (U.S. Patent 6,477,491) disclose a system that assigns a unique recording channel for each user at a meeting. Sharman et al. (U.S. Patent 6,100,882) disclose a system for creating a textual transcript. Wang (U.S. Patent 5,596,679) discloses a voting system for decoding speech in parallel with multiple speech recognizers. Bennett et al. (U.S. Patent 6,282,510) discloses a system that synchronizes the display of a recorded transcript to a plurality of computers.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L Albertalli whose telephone number is (703) 305-1817. The examiner can normally be reached on Mon - Fri, 8:00 AM - 5:30 PM, every second Fri off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 305-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

BLA 2/3/05

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER